



Seduta del

Sitzung vom

3.12.2025

DELIBERAZIONE N.

BESCHLUSS NR.

234

Oggetto:

Betreff:

Approvazione delle “Linee guida per l’utilizzo dell’intelligenza artificiale nella Regione Autonoma Trentino-Alto Adige/Südtirol” e istituzione del sistema di governance e trasparenza ai sensi della normativa vigente.

Genehmigung der „Leitlinien für den Einsatz künstlicher Intelligenz bei der Autonomen Region Trentino-Südtirol“ und Einrichtung eines Governance- und Transparenzsystems gemäß den geltenden Bestimmungen

Arno Kompatscher	Presidente/Präsident	presente/anwesend
Giulia Zanotelli	Vice Presidente sostituta del Presidente / Vizepräsidentin-Stellvertreterin des Präsidenten	presente/anwesend
Franz Thomas Locher	Vice Presidente/Vizepräsident	presente/anwesend
Carlo Daldoss	Assessore/Assessor	presente/anwesend
Angelo Gennaccaro	Assessore/Assessor	presente/anwesend
Luca Guglielmi	Assessore/Assessor	presente/anwesend
Gabriele Morandell	Segretaria generale della Giunta regionale / Generalsekretärin der Regionalregierung	presente/anwesend

Su proposta della Vice Presidente sostituta
del Presidente Giulia Zanotelli

Auf Vorschlag der Vizepräsidentin –
Stellvertreterin des Präsidenten Giulia Zanotelli

Ripartizione IV - Risorse strumentali

Abteilung IV – Vermögen und Beschaffung von
Gütern und Dienstleistungen

Ufficio informatica e digitalizzazione

Amt für Informatik und Digitalisierung

In riferimento all'oggetto la Giunta regionale ha discusso e deliberato quanto segue:

PREMESSO:

- che l'Intelligenza Artificiale (IA) è riconosciuta come una tecnologia a elevata capacità trasformativa in grado di incidere profondamente sui processi decisionali, produttivi e relazionali all'interno della Pubblica Amministrazione;
- che l'impiego dell'IA costituisce uno strumento strategico per migliorare l'efficienza dei servizi, semplificare le procedure amministrative e favorire un'interazione più immediata e trasparente tra amministrazione, cittadini e imprese;
- che l'introduzione dell'IA nei contesti lavorativi deve avvenire in maniera responsabile e regolamentata, rispettando i principi di trasparenza, tracciabilità, tutela dei diritti dei lavoratori e garantendo l'uso etico e la centralità dell'essere umano;
- che l'intero quadro regolatorio è ancorato ai principi costituzionali di legalità, imparzialità, buon andamento e trasparenza, che costituiscono la base valoriale imprescindibile per l'applicazione dell'IA in ambito pubblico.

VISTI:

- il Regolamento (UE) 2024/1689 (c.d. "AI Act"), che istituisce un sistema giuridico uniforme basato sulla classificazione del rischio, prevedendo requisiti stringenti per i sistemi ad alto rischio, l'obbligo di supervisione umana efficace, e la tenuta di un registro degli strumenti di IA impiegati;
- il Regolamento (UE) 2016/679 (c.d. "GDPR"), che stabilisce il quadro di

Die Regionalregierung hat über die oben genannte Angelegenheit beraten und Folgendes beschlossen:

VORAUSGESCHICKT:

- dass die Künstliche Intelligenz (KI) als eine Technologie mit hoher transformativer Fähigkeit gilt, die sich maßgeblich auf Entscheidungs-, Produktions- und Beziehungsprozesse innerhalb der öffentlichen Verwaltung auswirken kann;
- dass der Einsatz von KI ein strategisches Instrument darstellt, um die Effizienz der Dienste zu verbessern, Verwaltungsabläufe zu vereinfachen und eine unmittelbarere sowie transparentere Interaktion zwischen Verwaltung, Bürgern und Bürgerinnen sowie Unternehmen zu fördern;
- dass die Einführung von KI im Arbeitskontext verantwortungsbewusst und reguliert erfolgen muss, wobei die Grundsätze der Transparenz, der Rückverfolgbarkeit, des Schutzes der Rechte der Arbeitnehmenden einzuhalten sowie eine ethische Nutzung und die Zentralität des Menschen zu gewährleisten sind;
- dass der gesamte Rechtsrahmen an den Verfassungsgrundsätzen der Rechtmäßigkeit, Unparteilichkeit, guten Verwaltungsführung und Transparenz ausgerichtet ist, welche die unabdingbare Wertgrundlage für den Einsatz von KI im öffentlichen Bereich darstellen.

AUFGRUND:

- der Verordnung (EU) 2024/1689 (sog. „AI Act“), die ein einheitliches, auf Risikoeinstufung basierendes Rechtssystem schafft, in dem strengere Anforderungen für Hochrisiko-Systeme vorgesehen und eine wirksame menschliche Aufsicht sowie die Führung eines Registers der eingesetzten KI-Tools vorgeschrieben sind;
- der Verordnung (EU) 2016/679 (sog.

riferimento per la protezione dei dati personali, imponendo principi quali liceità, correttezza, trasparenza e minimizzazione dei dati;

- il Decreto Legislativo n. 82/2005 (c.d. “Codice dell’Amministrazione Digitale”), che definisce i diritti digitali dei cittadini e gli obblighi delle pubbliche amministrazioni;
- la Legge 132/2025, entrata in vigore il 10 ottobre 2025, che all’articolo 14 introduce l’obbligo di istituire un sistema di governance e trasparenza per l’uso dell’IA negli enti pubblici.

CONSIDERATO:

- che la Regione Autonoma Trentino-Alto Adige/Südtirol, nell’ambito delle proprie competenze in materia organizzativa, riconosce la necessità di definire un quadro organico di indirizzi e regole uniformi e vincolanti per l’adozione, lo sviluppo e la gestione dei sistemi di IA tra le diverse articolazioni dell’amministrazione, al fine di prevenire usi non controllati o disomogenei della tecnologia, nella formulazione di Linee guida regolamentari;
- che tali Linee guida mirano a garantire un approccio strutturato che integri valutazioni etiche, procedurali e organizzative, definendo l’IA come uno strumento al servizio dell’innovazione organizzativa e della qualità dei servizi pubblici;
- che le Linee guida prevedono l’applicazione dei principi di:
 - equità e non discriminazione (prevenendo bias algoritmico e garantendo parità di trattamento);
 - sicurezza e protezione dei dati (vietando l’introduzione di dati sensibili senza adeguate misure di sicurezza e

„DSGVO“), die den Rahmen für den Schutz personenbezogener Daten festlegt und Grundsätze wie Rechtmäßigkeit, Korrektheit, Transparenz und Datenminimierung vorsieht;

- des gesetzesvertretenden Dekrets Nr. 82/2005 (sog. „Kodex der digitalen Verwaltung“), der die digitalen Rechte der Bürger und Bürgerinnen und die Pflichten der öffentlichen Verwaltungen festlegt;
- des am 10. Oktober 2025 in Kraft getretenen Gesetzes Nr. 132/2025, das im Art. 14 die Pflicht zur Einrichtung eines Governance- und Transparenzsystems für den Einsatz von KI in öffentlichen Körperschaften vorsieht.

IN ANBETRACHT DER TATSACHE, DASS:

- die Autonome Region Trentino-Südtirol es im Rahmen ihrer Zuständigkeiten im organisatorischen Bereich für notwendig erachtet, durch die Festlegung entsprechender Leitlinien einen umfassenden Rahmen verbindlicher und einheitlicher Richtlinien und Regeln für die Einführung, Entwicklung und Verwaltung von KI-Systemen in den verschiedenen Verwaltungsbereichen zu definieren, um eine unkontrollierte oder uneinheitliche Nutzung der Technologie zu verhindern;
- diese Leitlinien darauf abzielen, einen strukturierten Ansatz zu gewährleisten, der ethische, verfahrenstechnische und organisatorische Bewertungen beinhaltet, und KI als Instrument zur Innovation von Organisation und Qualität der öffentlichen Dienste definiert;
- die Leitlinien die Anwendung folgender Grundsätze vorsehen:
 - Fairness und Nichtdiskriminierung (indem algorithmische Bias verhindert und Gleichbehandlung gewährleistet werden);
 - Sicherheit und Datenschutz (indem die Eingabe sensibler Daten ohne angemessene Sicherheits- und

- conformità al GDPR);
 - proporzionalità (commisurando l'uso dell'IA agli scopi istituzionali);
 - uso etico e centralità dell'essere umano (garantendo la supervisione umana e mantenendo l'utente titolare del processo e responsabile della decisione finale);
 - trasparenza e tracciabilità (dichiarando chiaramente quando un output è stato generato o integrato tramite IA);
 - accuratezza, verifica e validazione (sottoponendo ogni output a *checklist* e, nei documenti a maggiore impatto, a revisioni incrociate);
 - tutela dei lavoratori (nel rispetto del dialogo sociale e vietando forme di sorveglianza invasiva o controllo eccessivo).
 - che per l'attuazione e la supervisione delle Linee guida, è prevista l'istituzione di un organo collegiale per la governance, il coordinamento e il monitoraggio dei sistemi IA.
- DSGVO-Konformitätsmaßnahmen verboten wird);
 - Verhältnismäßigkeit (indem der KI-Einsatz auf die institutionellen Ziele abgestimmt wird);
 - ethische Nutzung und Zentralität des Menschen (indem die menschliche Aufsicht gewährleistet und die Nutzerverantwortung über den Prozess und die endgültige Entscheidung erhalten werden);
 - Transparenz und Rückverfolgbarkeit (indem ausdrücklich erklärt wird, wann ein Output durch KI generiert oder ergänzt wurde);
 - Genauigkeit, Überprüfung und Validierung (indem jeder Output anhand einer Checkliste überprüft und Dokumente mit größerer Auswirkung gegengeprüft werden);
 - Schutz der Arbeitnehmenden (indem der soziale Dialog eingehalten und invasive Überwachung oder übermäßige Kontrolle verboten werden);
 - zur Umsetzung und Überwachung der Leitlinien die Einrichtung eines Gremiums für Governance, Koordinierung und Monitoring der KI-Systeme vorgesehen ist;

Ritenuto di procedere all'approvazione delle "Linee guida per l'utilizzo dell'intelligenza artificiale nella Regione Autonoma Trentino-Alto Adige/Südtirol", in allegato al presente provvedimento quale parte integrante e sostanziale sotto la lettera "A";

Dato atto che le medesime Linee guida saranno adottate con un Decreto presidenziale per la loro entrata in vigore il 15° giorno successivo alla pubblicazione sul Bollettino ufficiale.

Vista la legge regionale 15 luglio 2009, n.3 "Norme in materia di bilancio e contabilità della Regione" e s.m., nonché il decreto legislativo 23 giugno 2011, n. 118 e s.m.

Ad unanimità di voti legalmente espressi,

Nach Dafürhalten, die „Leitlinien für den Einsatz künstlicher Intelligenz bei der Autonomen Region Trentino-Südtirol“, die unter Buchst. „A“ diesem Beschluss als ergänzender und wesentlicher Bestandteil beiliegen, zu genehmigen;

Nach Bestätigung der Tatsache, dass diese Leitlinien mit Dekret des Präsidenten der Region erlassen werden und am 15. Tag nach der Veröffentlichung im Amtsblatt in Kraft treten;

Aufgrund des Regionalgesetzes vom 15. Juli 2009, Nr. 3 „Bestimmungen über den Haushalt und das Rechnungswesen der Region“ i.d.g.F. sowie des gesetzesvertretenden Dekrets vom 23. Juni 2011, Nr. 118 i.d.g.F.;

delibera

1. di approvare le “Linee guida per l'utilizzo dell'intelligenza artificiale nella Regione Autonoma Trentino-Alto Adige/Südtirol”, allegate al presente atto per farne parte sostanziale e integrante sotto la lettera “A”;
2. di dare atto che le Linee guida di cui alla precedente punto 1. rappresentano il complesso delle regole uniformi e vincolanti nell'impiego dell'Intelligenza Artificiale per tutto il personale della Regione Autonoma Trentino-Alto Adige/Südtirol, compresi collaboratori esterni, consulenti e soggetti che operano in nome e per conto della Regione;
3. di istituire il “Tavolo tecnico regionale per l'Intelligenza Artificiale”, organo preposto alla governance, il coordinamento, l'adozione, l'implementazione ed il monitoraggio dei sistemi IA;
4. di dare atto che il Tavolo tecnico di cui al precedente punto 3. è presieduto dalla Segretaria Generale della Regione Autonoma Trentino-Alto Adige/Südtirol ed è composto dalla Dirigenza regionale, dal Data Protection Officer (DPO) dal Direttore dell'Ufficio Informatica e Digitalizzazione; il personale di questo Ufficio assicura l'attività di Segreteria del Tavolo in parola;
5. di dare atto che il Tavolo tecnico di cui al precedente punto 3. ha competenza per l'istruttoria per l'aggiornamento delle regole e delle procedure, la gestione delle anomalie e la valutazione e approvazione delle eventuali deroghe (eccezioni) all'applicazione delle Linee guida;
6. di istituire presso l'Ufficio Informatica e digitalizzazione il “Registro regionale degli strumenti di IA”, per le finalità delle Linee guida di cui al

beschließt die Regionalregierung

mit Einhelligkeit gesetzmäßig abgegebener Stimmen,

1. die „Leitlinien für den Einsatz künstlicher Intelligenz bei der Autonomen Region Trentino-Südtirol“, die unter Buchst. „A“ diesem Beschluss als ergänzender und wesentlicher Bestandteil beiliegen, zu genehmigen;
2. zu bestätigen, dass die Leitlinien laut Z. 1 das Regelwerk der verbindlichen und einheitlichen Bestimmungen für den Einsatz künstlicher Intelligenz für das gesamte Personal der Autonomen Region Trentino-Südtirol, einschließlich externer Mitarbeiter/Mitarbeiterinnen, Berater/Beraterinnen und Personen, die im Namen und Auftrag der Region tätig sind, darstellen;
3. die „Technische Arbeitsgruppe der Region für Künstliche Intelligenz“ als Gremium einzurichten, das für Governance, Koordinierung, Einführung, Implementierung und Monitoring der KI-Systeme zuständig ist;
4. zu bestätigen, dass die Generalsekretärin der Autonomen Region Trentino-Südtirol den Vorsitz der Technischen Arbeitsgruppe laut Z. 3 führt, die sich aus den Führungskräften, dem Data Protection Officer (DPO) und dem Direktor des Amtes für Informatik und Digitalisierung zusammensetzt, wobei das Personal dieses Amtes die Schriftführung für die Arbeitsgruppe gewährleistet;
5. zu bestätigen, dass die Technische Arbeitsgruppe laut Z. 3 für die Ausarbeitung und die Aktualisierung der Bestimmungen und Verfahren, für die Verwaltung von Anomalien und die Bewertung sowie Genehmigung etwaiger Ausnahmen zur Anwendung der Leitlinien zuständig ist;
6. beim Amt für Informatik und

precedente punto 1, il quale opera secondo le direttive del Tavolo tecnico di cui al precedente punto 3;

7. di stabilire che, alla data di entrata in vigore delle presenti Linee guida, gli strumenti di IA iscritti nel Registro e come tali acconsentiti per l'uso tramite gli account aziendali (@regione.taa.it) sono:
- Gemini IA (assistente basato sui modelli linguistici Gemini);
 - Gemini Workspace (IA integrata in Gmail, Documenti, Fogli, Presentazioni e Drive).

Digitalisierung das „Regionale Register der KI-Tools“ für die Zwecke der Leitlinien laut Z. 1 einzurichten, das gemäß den Vorgaben der Arbeitsgruppe laut vorstehender Z. 3 geführt wird;

7. festzulegen, dass mit Inkrafttreten dieser Leitlinien nachstehende KI-Tools im Register eingetragen und somit für die Nutzung über Accounts der Verwaltung (@regione.taa.it) zugelassen sind:
- Gemini IA (auf den Gemini-Sprachmodellen basierender Assistent);
 - Gemini Workspace (in Gmail, Docs, Sheets, Präsentationen und Drive integrierte KI).

Contro il presente provvedimento sono ammessi alternativamente i seguenti ricorsi:

- ricorso giurisdizionale al TRGA di Trento ai sensi dell'art. 29 del D.Lgs. 2 luglio 2010 n. 104;
- ricorso straordinario al Presidente della Repubblica da parte di chi vi abbia interesse entro 120 giorni ai sensi del DPR 24.11.1971 n. 1199.

Gegen diese Maßnahme können alternativ nachstehende Rekurse eingelegt werden:

- Rekurs beim Regionalen Verwaltungsgericht Trient im Sinne des Art. 29 des GvD vom 2. Juli 2010, Nr. 104;
- außerordentlicher Rekurs an den Präsidenten der Republik, der von Personen, die ein rechtliches Interesse daran haben, binnen 120 Tagen im Sinne des DPR vom 24.11.1971, Nr. 1199 einzulegen ist.

Letto, confermato e sottoscritto.

IL PRESIDENTE

DER PRÄSIDENT

Arno Kompatscher
firmato digitalmente / digital signiert

Gelesen, bestätigt und unterzeichnet

LA SEGRETARIA GENERALE
DELLA GIUNTA REGIONALE

DIE GENERALSEKRETÄRIN DER
REGIONALREGIERUNG

Gabriele Morandell
firmato digitalmente / digital signiert

Questo documento, se trasmesso in forma cartacea, costituisce copia dell'originale informatico firmato digitalmente, valido a tutti gli effetti di legge, predisposto e conservato presso questa Amministrazione (D.Lgs 82/05). L'indicazione del nome del firmatario sostituisce la sua firma autografa (art. 3 D. Lgs. 39/93).

Falls dieses Dokument in Papierform übermittelt wird, stellt es eine für alle gesetzlichen Wirkungen gültige Kopie des elektronischen digital signierten Originals dar, das von dieser Verwaltung erstellt und bei derselben aufbewahrt wird (GvD Nr. 82/2005). Die Angabe des Namens der unterzeichnenden Person ersetzt deren eigenhändige Unterschrift (Art. 3 GvD Nr. 39/1993).

A

“Linee guida per l’utilizzo dell’intelligenza artificiale nella Regione Autonoma Trentino-Alto Adige/Südtirol”

Sommario	
Introduzione	3
1 Riferimenti normativi a livello europeo e nazionale	4
1.1 Principi ispiratori e finalità della normativa europea e nazionale	5
1.1.1 Equità e non discriminazione	6
1.1.2 Sicurezza e protezione dei dati	6
1.1.3 Proporzionalità	7
1.1.4 Uso etico e centralità dell'essere umano	7
1.1.5 Trasparenza e tracciabilità	7
1.1.6 Accuratezza, verifica e validazione	7
1.1.7 Etica e non discriminazione	8
1.1.8 Tutela dei lavoratori	8
2. Le “Linee guida per l’utilizzo dell’intelligenza artificiale nella Regione Autonoma Trentino-Alto Adige/Südtirol”	8
2.1 Obiettivi	8
2.2 Ambito di applicazione e governance	9
2.2.1 Destinatari	9
2.2.2 Tavolo tecnico regionale per l’Intelligenza Artificiale	9
2.2.3 Esclusioni	10
2.2.4 Gestione delle eccezioni	10
2.3 Registro, modalità operative e responsabilità	11
2.3.1 Registro regionale degli strumenti IA	11
2.3.2 Linee di condotta per il personale e prima soluzione di IA consentita	11
2.4 Esempi pratici di utilizzo e casi d’uso	13
2.4.1 Uso consentito	13
2.4.2 Uso non consentito	15
2.5. Formazione obbligatoria	16
2.6. Monitoraggio e revisione	16
2.6.1 Indicatori di monitoraggio	16
2.6.2 Revisione periodica	17
Glossario	18

Introduzione

L'Intelligenza Artificiale (IA) si configura oggi come una tecnologia a elevata capacità trasformativa, in grado di incidere profondamente sui processi decisionali, produttivi e relazionali all'interno della Pubblica Amministrazione. Il suo impiego non rappresenta semplicemente un'opportunità tecnologica, ma costituisce uno strumento strategico per migliorare l'efficienza dei servizi, semplificare le procedure amministrative e favorire un'interazione più immediata e trasparente tra amministrazione, cittadini e imprese.

L'introduzione dell'IA nei contesti lavorativi deve avvenire in maniera responsabile e regolamentata. È essenziale che ogni sistema basato su IA sia implementato nel rispetto dei principi di trasparenza, tracciabilità e tutela dei diritti dei lavoratori, garantendo al contempo un accesso equo e inclusivo alle tecnologie emergenti.

L'adozione consapevole dell'IA richiede pertanto un approccio strutturato, che integri valutazioni etiche, procedurali e organizzative, affinché gli effetti sul lavoro, sulle dinamiche interne agli uffici ed i possibili effetti nel contesto istituzionale e sociale siano positivi e sostenibili.

In questa prospettiva, l'IA non è un fine in sé, ma uno strumento al servizio dell'innovazione organizzativa e della qualità dei servizi pubblici. La sua introduzione deve essere accompagnata da misure di governance chiare, dalla formazione del personale e da un monitoraggio continuo dei processi decisionali automatizzati, in conformità alle normative nazionali ed europee vigenti.

In tal senso, l'Intelligenza Artificiale sta già modificando in modo sostanziale il panorama lavorativo, esercitando effetti differenziati sui vari settori economici e produttivi. Nei processi amministrativi, l'IA consente di automatizzare attività ripetitive, ottimizzare l'impiego delle risorse e supportare le decisioni strategiche mediante l'analisi avanzata dei dati e la previsione di tendenze laddove opportune e necessarie rispetto alla mission dell'organizzazione. Tecnologie basate su IA, quali sistemi di apprendimento automatico, chatbot, assistenti virtuali e piattaforme di

analisi predittiva, contribuiscono a incrementare l'efficienza operativa, consentendo alle organizzazioni di affrontare sfide complesse con soluzioni innovative.

1 Riferimenti normativi a livello europeo e nazionale

A livello europeo, il quadro normativo è delineato principalmente dal **Regolamento (UE) 2024/1689**, noto come "AIAct", che istituisce un sistema giuridico uniforme per l'intero territorio dell'Unione. Tale regolamento introduce un approccio basato sulla classificazione del rischio, raggruppando i sistemi di IA in quattro categorie di rischio: inaccettabile, alto, limitato e minimo. I sistemi a rischio inaccettabile, come quelli di social scoring o di manipolazione comportamentale, sono vietati in quanto incompatibili con i valori fondamentali dell'Unione. I sistemi ad alto rischio, che includono applicazioni in ambiti sensibili, quali la sicurezza e la gestione del personale, sono ammessi solo a condizione che rispettino requisiti stringenti di sicurezza, trasparenza e controllo umano. I sistemi a rischio limitato, come chatbot e strumenti di generazione di contenuti sintetici o deepfake, sono soggetti a specifici obblighi di trasparenza nei confronti degli utenti, mentre i sistemi a rischio minimo possono essere utilizzati liberamente nel rispetto dei principi generali del regolamento.

Tra gli obblighi principali previsti dall'"AI Act" rientrano la valutazione d'impatto sui diritti fondamentali e sui rischi sistemici connessi all'uso dell'IA, la tenuta di un registro degli strumenti di IA impiegati, l'obbligo di garantire una supervisione umana efficace in tutte le fasi del processo decisionale automatizzato e la trasparenza informativa nei confronti degli utenti. Tali misure intendono assicurare la tracciabilità delle decisioni, la verificabilità dei risultati e la prevenzione di effetti discriminatori o lesivi dei diritti individuali.

Parallelamente, il **Regolamento (UE) 2016/679**, comunemente noto come "General Data Protection Regulation" (GDPR), stabilisce il quadro di riferimento per la protezione dei dati personali e introduce i principi di liceità, correttezza, trasparenza, minimizzazione dei dati, limitazione della conservazione, integrità e riservatezza. Il GDPR riconosce ai soggetti interessati diritti fondamentali, tra cui l'accesso, la

rettifica, la cancellazione, la limitazione del trattamento, l'opposizione e la portabilità dei propri dati. La figura del Responsabile della Protezione dei Dati (Data Protection Officer – DPO) assume un ruolo centrale, in quanto incaricata di fornire consulenza, monitorare l'applicazione della normativa, promuovere attività di formazione interna e fungere da punto di contatto con l'Autorità Garante per la protezione dei dati personali.

Sul piano nazionale, il “Codice dell'Amministrazione Digitale” (**D.Lgs. n. 82/2005**) definisce i diritti digitali dei cittadini e gli obblighi delle pubbliche amministrazioni in materia di gestione dei servizi online, identità digitale, interoperabilità dei sistemi, sicurezza delle infrastrutture, conservazione e accessibilità dei documenti informatici. In stretta correlazione, la normativa sulla trasparenza e sulla prevenzione della corruzione (D.Lgs. n. 33/2013) impone obblighi di pubblicità e tracciabilità delle attività amministrative, con l'obiettivo di garantire un uso dell'IA conforme ai principi di imparzialità e buon andamento dell'azione pubblica.

Un ruolo di rilievo è svolto anche dalle linee guida e dai documenti tecnici elaborati da organismi nazionali e comunitari. L'Agenzia per l'Italia Digitale (AgID) definisce standard e raccomandazioni in materia di interoperabilità, sicurezza, adozione del cloud e gestione documentale; l'ENISA, Agenzia dell'Unione Europea per la cybersicurezza, fornisce indicazioni operative sulla gestione dei rischi informatici e sulla sicurezza dei sistemi digitali; il Garante per la protezione dei dati personali pubblica pareri, raccomandazioni e provvedimenti sanzionatori per assicurare il rispetto della normativa in materia di privacy e trattamento dei dati.

Infine, la recente **legge 132/2025** (entrata in vigore il 10 ottobre 2025) all'articolo 14, ("Uso dell'intelligenza artificiale nella pubblica amministrazione"), introduce l'obbligo di istituire un sistema di governance e trasparenza per l'uso dell'IA negli enti pubblici.

1.1 Principi ispiratori e finalità della normativa europea e nazionale

L'intero quadro regolatorio è ancorato ai principi costituzionali di legalità, imparzialità, buon andamento e trasparenza, che rappresentano i cardini dell'azione amministrativa e costituiscono la base valoriale imprescindibile per l'applicazione

dell'Intelligenza Artificiale in ambito pubblico. Questi principi assicurano che l'innovazione tecnologica non comprometta le garanzie democratiche, ma si sviluppi come strumento di supporto al servizio della collettività, nel rispetto della dignità e dei diritti delle persone. In particolare, nell'utilizzo dell'IA l'Amministrazione regionale, i suoi dipendenti ed i soggetti che operano in nome per conto della medesima Amministrazione sono tenuti al rispetto dei seguenti principi.

1.1.1 Equità e non discriminazione

L'adozione di sistemi di Intelligenza Artificiale deve essere improntata ai principi di equità, imparzialità e uguaglianza di trattamento. È necessario prevenire e correggere qualsiasi forma di bias algoritmico che possa determinare discriminazioni dirette o indirette basate su genere, etnia, età, orientamento, condizione socioeconomica o altre caratteristiche personali. A tal fine, vanno attuati processi di monitoraggio continuo dei risultati generati dai sistemi, con l'obiettivo di individuare eventuali deviazioni sistematiche e intervenire con misure correttive. L'equità deve essere garantita anche nella fase di progettazione e addestramento dei modelli, mediante la selezione di dataset rappresentativi e la validazione indipendente degli algoritmi. La trasparenza dei criteri di valutazione e la possibilità di revisione umana degli output costituiscono ulteriori garanzie di imparzialità.

1.1.2 Sicurezza e protezione dei dati

La sicurezza informatica e la protezione dei dati personali rappresentano prerequisiti essenziali per l'impiego dell'Intelligenza Artificiale. È vietato introdurre nei sistemi di IA dati sensibili, classificati o comunque idonei a rivelare informazioni riservate, salvo specifiche autorizzazioni e previa adozione di misure di sicurezza adeguate, come la cifratura, la pseudonimizzazione o l'accesso controllato. I trattamenti devono essere conformi ai principi stabiliti dal GDPR e dalle linee guida del Garante per la protezione dei dati personali, assicurando che i dati siano raccolti per finalità determinate, pertinenti e limitate allo scopo istituzionale. Parallelamente all'impiego degli strumenti di Intelligenza Artificiale devono inoltre essere implementate politiche di data governance che garantiscano l'integrità, la disponibilità e la tracciabilità dei dati, con piani di risposta agli incidenti (incidentresponse) e sistemi di audit periodico per prevenire violazioni o accessi non autorizzati.

1.1.3 Proporzionalità

Il principio di proporzionalità richiede che l'uso dell'Intelligenza Artificiale sia strettamente commisurato agli scopi istituzionali e alle finalità del trattamento dei dati. È necessario evitare un impiego eccessivo o ridondante della tecnologia, così come la raccolta di informazioni non necessarie o sproporzionate rispetto agli obiettivi perseguiti. Deve essere valutato l'impatto dell'IA non solo in termini di efficienza operativa, ma anche in relazione ai potenziali effetti sulla privacy, sull'autonomia decisionale dei lavoratori e sulla qualità del servizio offerto ai cittadini. La proporzionalità si traduce quindi in un equilibrio dinamico tra innovazione e tutela, garantendo che l'automazione non sostituisca ma supporti le competenze umane.

1.1.4 Uso etico e centralità dell'essere umano

I sistemi di Intelligenza Artificiale devono essere considerati strumenti ausiliari, progettati con l'obiettivo di migliorare il benessere collettivo e la qualità dei servizi pubblici, senza sostituire le capacità intellettuali, creative e decisionali degli individui. La supervisione umana rimane un principio irrinunciabile: l'utente è sempre titolare del processo e mantiene la responsabilità della decisione finale che non può essere delegata integralmente alla macchina. L'IA, in questo contesto, svolge una funzione di supporto e di potenziamento delle capacità umane, contribuendo a ottimizzare l'efficienza operativa e la gestione delle informazioni, ma senza assumere autonomia decisionale in merito a scelte sensibili o discrezionali.

1.1.5 Trasparenza e tracciabilità

L'impiego di strumenti di IA nelle attività lavorative deve essere caratterizzato da trasparenza e tracciabilità. È essenziale dichiarare in modo chiaro quando un output, sia esso un documento, una nota o una risposta a interrogazioni, sia stato generato o integrato tramite un sistema di IA. Questo principio consente di preservare la fiducia nei risultati prodotti e di distinguere l'algoritmo come strumento operativo e non come autore delle informazioni.

1.1.6 Accuratezza, verifica e validazione

Ogni output prodotto da un sistema di IA deve essere sottoposto a un processo rigoroso di verifica e validazione, volto a garantire accuratezza, completezza e pertinenza delle informazioni. La valutazione critica è indispensabile poiché i sistemi di IA possono generare risultati incompleti, datati o, in alcuni casi, "allucinazioni",

ossia informazioni false o fuorvianti. Per mitigare tali rischi, è necessario adottare strumenti operativi quali checklist di validazione per ogni output e revisioni incrociate da parte di almeno due operatori nei documenti a maggiore impatto. Questo approccio garantisce un controllo qualitativo sistematico, riducendo la probabilità di errori e aumentando la fiducia nell'utilizzo dell'IA nei processi decisionali e operativi.

1.1.7 Etica e non discriminazione

I sistemi di IA devono essere progettati e utilizzati in modo da prevenire bias e discriminazioni, rispettando i principi etici fondamentali e promuovendo attivamente l'inclusività. L'obiettivo è assicurare un trattamento equo e imparziale per tutti gli individui e i gruppi, con particolare attenzione alle categorie più vulnerabili. È necessario monitorare e correggere eventuali distorsioni presenti nei dati o negli algoritmi, adottando strategie di mitigazione per ridurre l'impatto dei bias e garantire che i risultati generati siano equi, trasparenti e conformi ai valori etici e normativi della Pubblica Amministrazione.

1.1.8 Tutela dei lavoratori

L'introduzione di sistemi di Intelligenza Artificiale nei luoghi di lavoro deve avvenire nel rispetto del dialogo sociale e delle normative sul lavoro. L'utilizzo di strumenti intelligenti per monitorare le attività dei dipendenti richiede regole precise per evitare forme di sorveglianza invasiva o controllo eccessivo, che potrebbero compromettere la libertà, la dignità e la qualità del lavoro. La tecnologia deve supportare l'efficienza organizzativa e la sicurezza dei processi senza incidere negativamente sulle condizioni lavorative, garantendo un equilibrio tra innovazione, produttività e tutela dei diritti dei lavoratori.

2. Le “Linee guida per l'utilizzo dell'intelligenza artificiale nella Regione Autonoma Trentino-Alto Adige/Südtirol”

2.1 Obiettivi

Nel quadro dei principi generali indicati al precedente capitolo 1, la Regione Autonoma Trentino-Alto Adige/Südtirol, anche in virtù delle prerogative conferite dallo Statuto speciale e nell'ambito delle proprie competenze di governo territoriale, riconosce la necessità di definire, nel contesto e a completamento della già

richiamata normativa europea e nazionale, un quadro organico di indirizzi e regole uniformi e vincolanti (Linee guida) per l'adozione, lo sviluppo e la gestione dei sistemi di Intelligenza Artificiale tra le diverse articolazioni dell'amministrazione, per prevenire usi non controllati o disomogenei della tecnologia. Tale iniziativa si inserisce nel processo di trasformazione digitale del settore pubblico e risponde all'esigenza di garantire un impiego dell'IA conforme ai principi generali che la riguardano.

2.2 Ambito di applicazione e governance

2.2.1 Destinatari

Le Linee guida trovano applicazione nei confronti di tutto il personale della Regione Autonoma Trentino-Alto Adige/Südtirol, compresi i collaboratori esterni, i consulenti e i soggetti che, a qualsiasi titolo, utilizzano sistemi di Intelligenza Artificiale per conto della Regione o delle sue articolazioni operative. L'ambito di applicazione si estende anche ad società, enti ed organismi controllati o coordinati dall'amministrazione regionale, nei limiti delle rispettive competenze e funzioni.

Rientrano inoltre nel perimetro applicativo tutte le attività che implicano l'uso di account, strumenti digitali o basi di dati acquisiti o gestiti nell'ambito delle funzioni lavorative, indipendentemente dal tipo di dispositivo utilizzato, sia esso fornito dall'amministrazione o di proprietà del dipendente. Ciò include ogni forma di interazione con sistemi di IA finalizzata all'elaborazione di dati, alla produzione di contenuti o al supporto decisionale in ambito istituzionale.

2.2.2 Tavolo tecnico regionale per l'Intelligenza Artificiale

È istituito il "Tavolo tecnico regionale per l'Intelligenza Artificiale" per la governance, il coordinamento, l'adozione, l'implementazione ed il monitoraggio di sistemi IA. La composizione del Tavolo è rimessa ad uno specifico provvedimento giuntale e deve includere la dirigenza regionale, personale incaricato dell'Ufficio Informatica e digitalizzazione, il Data Protection Officer (DPO) e può essere integrato, di volta in volta, rispetto alle tematiche in esame, con altro personale competente per materia. Al Tavolo tecnico è assegnata l'istruttoria per l'aggiornamento delle regole e delle procedure in funzione dell'evoluzione normativa e tecnologica, ivi comprese presenti Linee guida, nonché la valutazione delle segnalazioni e la gestione delle anomalie, incluse sospensioni temporanee, audit interni o revisione delle policy.

Il Tavolo tecnico esercita altresì le competenze di cui al successivo paragrafo 2.2.4.

2.2.3 Esclusioni

Sono escluse dall'ambito di applicazione delle presenti Linee guida le tecnologie informatiche che non presentano caratteristiche riconducibili all'Intelligenza Artificiale, ossia quegli strumenti che operano secondo logiche deterministiche e non basate su apprendimento o adattamento autonomo. Rientrano tra le esclusioni anche le automazioni ordinarie prive di capacità autonoma di apprendimento o di generazione di decisioni, come ad esempio i sistemi di workflow statici, le macro-applicative o i software di gestione documentale che non incorporano moduli predittivi o generativi.

Tali esclusioni non esonerano tuttavia gli operatori dall'obbligo di garantire la sicurezza, la riservatezza e la corretta gestione dei dati trattati, che restano soggetti alle normative vigenti in materia di protezione delle informazioni e di buon andamento amministrativo.

2.2.4 Gestione delle eccezioni

Eventuali deroghe all'applicazione delle presenti Linee guida possono essere ammesse esclusivamente in casi motivati e circoscritti, previa valutazione e approvazione da parte del Tavolo tecnico di cui al precedente paragrafo 2.2.2. La richiesta di deroga deve essere accompagnata da una documentazione dettagliata che espliciti le ragioni dell'eccezione, la natura del sistema o del progetto interessato, nonché le misure di mitigazione dei rischi eventualmente adottate per garantire la conformità ai principi generali di legalità, trasparenza e sicurezza.

Il citato Tavolo tecnico ha il compito di esaminare le richieste, verificare la coerenza con il quadro normativo e valutare gli impatti potenziali sull'organizzazione e sui diritti degli utenti. Le decisioni assunte devono essere tracciate e conservate ai fini di controllo e revisione periodica, assicurando che le eccezioni restino eccezionali e non si traducano in una prassi operativa ordinaria.

In tal modo, la gestione delle eccezioni si configura come strumento di equilibrio tra l'esigenza di flessibilità amministrativa e la necessità di mantenere elevati standard di affidabilità, sicurezza e responsabilità nell'uso dell'Intelligenza Artificiale in ambito regionale.

2.3 Registro, modalità operative e responsabilità

2.3.1 Registro regionale degli strumenti IA

È istituito presso l'Ufficio Informatica e digitalizzazione il "Registro regionale degli strumenti di IA", che opera secondo le direttive del Tavolo tecnico di cui al precedente paragrafo 2.2.2, nella finalità di garantire la conformità tecnica, giuridica e organizzativa delle soluzioni impiegate

Il Registro contiene un elenco aggiornato degli strumenti di Intelligenza Artificiale autorizzati, corredato, se opportuno, da schede tecniche che ne descrivono le caratteristiche principali, le finalità di impiego e gli ambiti di utilizzo. Le schede includono anche informazioni relative alla valutazione di impatto, alle misure di sicurezza implementate, ai livelli di supervisione umana previsti e ad eventuali limitazioni operative. Ogni strumento registrato è monitorato rispetto allo stato di aggiornamento, alle revisioni effettuate e alle eventuali segnalazioni o sospensioni che ne possano compromettere l'affidabilità o la conformità normativa.

L'inserimento di un nuovo strumento di Intelligenza Artificiale nel Registro è subordinato a un processo di valutazione articolato su tre dimensioni:

1. La valutazione tecnica: riguarda la sicurezza informatica, l'interoperabilità con i sistemi esistenti e l'affidabilità operativa dello strumento, garantendo che l'applicativo possa essere impiegato senza compromettere l'integrità dei dati o l'efficienza dei processi.
2. La valutazione giuridica: verifica la conformità dello strumento alle normative vigenti, includendo la necessità di eventuali Data Protection Impact Assessment (DPIA) e l'aderenza ai requisiti dell'AI Act, del GDPR e delle altre disposizioni regolatorie pertinenti.
3. La valutazione organizzativa: analizza l'impatto dell'introduzione dello strumento sui processi interni, individuando le procedure necessarie, le modifiche operative e la formazione del personale richiesta per un utilizzo sicuro ed efficace.

2.3.2 Linee di condotta per il personale e prima soluzione di IA consentita

Ogni dipendente, collaboratore o consulente che faccia uso di sistemi di Intelligenza Artificiale per conto della Regione Autonoma Trentino-Alto Adige/Südtirol è tenuto –

si ribadisce - a operare secondo principi di diligenza, responsabilità e trasparenza, garantendo che l'utilizzo di tali strumenti avvenga nel pieno rispetto delle normative vigenti e delle disposizioni organizzative interne. L'impiego dell'IA deve essere sempre accompagnato da una supervisione umana attiva, volta a verificare la correttezza, l'affidabilità e la pertinenza degli output generati. Ogni risultato prodotto o supportato da un sistema di IA deve essere sottoposto a validazione da parte dell'operatore responsabile, che ne assume la piena responsabilità sotto il profilo tecnico, giuridico e contenutistico.

In particolare, il personale è tenuto a segnalare esplicitamente, in tutti i documenti ufficiali, relazioni, comunicazioni interne o atti amministrativi, quando il contenuto sia stato redatto, modificato o integrato con l'ausilio di un sistema di Intelligenza Artificiale, utilizzando in calce al documento la seguente frase o simile *“Istruito anche con l'ausilio di I.A”*. Tale obbligo di trasparenza risponde al principio di tracciabilità dei processi decisionali e garantisce la possibilità di ricostruire il ruolo svolto dall'IA nell'elaborazione del documento, in conformità ai requisiti di accountability e di controllo pubblico. Non è necessario indicare l'uso dell'IA per attività interne o per la generazione di bozze preliminari che vengono poi significativamente modificate dall'utente, ferma ovviamente la verifica e validazione finale da parte del medesimo.

È fatto divieto di delegare integralmente a sistemi di IA decisioni che comportino valutazioni discrezionali, specialmente in ambiti sensibili o suscettibili di produrre effetti diretti sui diritti e sugli interessi delle persone. Tra questi rientrano, a titolo esemplificativo, le decisioni relative alla concessione di benefici economici, all'irrogazione di sanzioni, alla gestione di procedimenti disciplinari o alla valutazione delle performance del personale. In tali casi, il ruolo dell'IA può essere esclusivamente di supporto informativo o analitico, mentre la decisione finale deve restare di esclusiva competenza umana.

Dato atto che l'utilizzo dei sistemi di Intelligenza Artificiale è consentito solo se tali strumenti risultano iscritti nel Registro regionale degli strumenti di IA, il personale deve verificare preventivamente che l'applicativo utilizzato sia autorizzato e aggiornato, nonché conforme ai requisiti di sicurezza, protezione dei dati e

interoperabilità. L'uso di strumenti non registrati o non conformi costituisce violazione delle presenti linee guida e può comportare responsabilità disciplinari o amministrative.

Gli strumenti di IA consentiti alla data di adozione delle presenti Linee guida, in quanto inseriti nel previsto Registro di cui al precedente paragrafo 2.3.1 e già attivi sugli account aziendali di posta elettronica (con dominio @regione.taa.it) sono i seguenti:

- **Gemini IA:** si tratta di un assistente basato sull'Intelligenza Artificiale che sfrutta i modelli linguistici Gemini di Google. Progettato per potenziare la creatività e la produttività degli utenti, offre una vasta gamma di funzionalità.
- **Gemini Workspace:** questa è l'intelligenza artificiale di Gemini integrata direttamente nella suite di applicazioni Google per la produttività e la collaborazione, quali Gmail, Documenti, Fogli, Presentazioni e Drive.

Non è in alcun modo consentito l'utilizzo di strumenti di IA (e di GenAI) diversi da quelli espressamente autorizzati ai sensi delle presenti Linee guida. In ogni caso, non è consentito l'utilizzo di strumenti basati su IA (o GenAI) diversi da quelli autorizzati ai sensi delle presenti Linee guida, per il tramite di account personali laddove siano utilizzati per finalità lavorative.

2.4 Esempi pratici di utilizzo e casi d'uso

2.4.1 Uso consentito

L'adozione dei sistemi di Intelligenza Artificiale all'interno della Regione Autonoma Trentino-Alto Adige/Südtirol comporta una serie di benefici trasversali che interessano la qualità del lavoro, l'accesso alle informazioni, la produttività, la creatività, la collaborazione, la qualità delle decisioni, la sicurezza e la sostenibilità delle attività amministrative. Tuttavia, deve essere sempre garantita la supervisione umana (human-in-the-loop) e la piena responsabilità in capo all'operatore. L'utilizzo deve rispettare i principi di legalità, trasparenza e non discriminazione e quanto declinato al precedente paragrafo 1.1 delle presenti Linee guida.

Gli ambiti di utilizzo consentito includono:

A. Supporto Operativo e Analisi Documentale

L'IA può essere impiegata per ottimizzare la gestione interna dei flussi di lavoro, riducendo le attività ripetitive. A titolo esemplificativo all'IA può essere rimesso il ruolo/la funzione di

- Assistente per la stesura di documenti non riservati: generazione di documenti complessi, discussioni o comunicazioni elettroniche, per la ricerca di fonti su grandi quantità di dati, per il brainstorming e individuare nuovi spunti per l'elaborazione della sintesi finale.
- Redazione di bozze e contenuti: supporto nella stesura di bozze iniziali di testi, relazioni, comunicazioni interne o documenti non vincolanti, purché l'output sia sottoposto a verifica e validazione rigorosa da parte del personale.
- Classificazione e smistamento documentale: utilizzo di sistemi IA per la classificazione automatica e lo smistamento di documenti in ingresso, con monitoraggio continuo delle anomalie e formazione specifica per il personale addetto.

B. Miglioramento dei Servizi al Cittadino

L'IA può potenziare l'accessibilità e la tempestività delle risposte al pubblico.

- Assistenza Virtuale (Chatbot): implementazione di chatbot e assistenti virtuali per fornire risposte immediate a domande frequenti (FAQ) e richieste di carattere generale da parte dei cittadini e del personale interno. Per risposte complesse o pubblicazioni ufficiali, è obbligatoria la validazione umana preventiva.
- Accessibilità e Inclusività: supporto linguistico multilingue e strumenti che abbattano barriere comunicative o fisiche, garantendo un servizio pubblico più equo e fruibile.

C. Supporto ai Processi Decisionali e Strategici

L'IA può fornire analisi avanzate, migliorando la qualità delle decisioni umane.

- Analisi Predittiva e Simulazioni: elaborazione di grandi volumi di dati per formulare stime, scenari futuri, reportistica avanzata e simulazioni che supportino le decisioni strategiche e gestionali.
- Analisi Preliminare per Contributi/Bandi: utilizzo di sistemi per l'analisi preliminare di domande di concessione di benefici economici o

finanziamenti, purché la decisione finale rimanga sempre affidata al funzionario umano. Devono essere previsti audit periodici sui risultati per prevenire bias.

In ogni caso, vale ribadire, è obbligatorio:

1. verificare preventivamente che lo strumento sia iscritto e autorizzato nel Registro regionale degli strumenti IA.
2. includere in tutti gli atti e documenti ufficiali una dichiarazione esplicita sull'utilizzo del supporto di Intelligenza Artificiale, come indicato al precedente paragrafo 2.3.2, secondo capoverso.
3. eseguire una Data Protection Impact Assessment (DPIA) prima dell'implementazione di sistemi che trattano dati personali, come richiesto dalla normativa.

2.4.2 Uso non consentito

L'impiego dei sistemi di Intelligenza Artificiale nella Regione Autonoma Trentino-Alto Adige/Südtirol è soggetto a limiti rigorosi volti a tutelare la sicurezza, la privacy, l'etica e la conformità normativa. È vietato inserire nei sistemi dati personali o informazioni classificate come riservate o confidenziali, al fine di prevenire rischi di esposizione indebita, accessi non autorizzati o violazioni della normativa sulla protezione dei dati.

Non è consentito utilizzare i sistemi di IA per generare o diffondere contenuti discriminatori, offensivi, violenti, illegali o che possano ledere i diritti di terzi, così come è proibito presentare l'output di un sistema di IA come lavoro originale senza dichiararne l'effettivo utilizzo, pena la violazione dei principi di trasparenza e integrità.

L'uso dei sistemi di IA per prendere decisioni automatiche di impatto significativo su individui o processi senza supervisione e validazione umana è vietato, in quanto compromette la responsabilità, la tracciabilità e il principio della centralità dell'essere umano nei processi decisionali. Allo stesso modo, non è ammesso impiegare l'IA per ottenere accesso non autorizzato a informazioni o sistemi, né per attività non lavorative o personali che possano interferire con le funzioni istituzionali dell'ente.

È inoltre vietato utilizzare tecnologie basate su riconoscimento emotivo, eye-tracking o analisi del sentiment per la sorveglianza dei lavoratori, in quanto tali strumenti

possono compromettere la privacy, il benessere e i diritti dei dipendenti. L'uso dei sistemi di IA deve sempre rispettare leggi, regolamenti e policy interne dell'ente, garantendo che le applicazioni siano coerenti con finalità istituzionali, etiche e operative.

2.5. Formazione obbligatoria

La competente Struttura regionale in tema di formazione del personale indica, nella prevista programmazione, i percorsi formativi periodici obbligatori destinati a tutto il personale coinvolto nell'uso dei sistemi di Intelligenza Artificiale. Tali programmi hanno l'obiettivo di sviluppare competenze multidisciplinari, combinando conoscenze tecniche, giuridiche ed etiche. I moduli trattano le potenzialità e i limiti dell'IA, le competenze tecniche necessarie per l'utilizzo sicuro e corretto dei sistemi, nonché gli aspetti normativi e di compliance legati alla protezione dei dati e alla supervisione dei processi automatizzati. Particolare attenzione è dedicata alla cultura etica dell'innovazione, alla rilevazione e mitigazione dei bias, alla prevenzione di discriminazioni, alla trasparenza e tracciabilità dei processi e alla definizione dei ruoli e responsabilità nella supervisione umana. I corsi includono inoltre analisi di casi d'uso concreti e scenari operativi, al fine di consolidare le competenze acquisite e favorire un'applicazione responsabile e consapevole dei sistemi IA.

2.6. Monitoraggio e revisione

2.6.1 Indicatori di monitoraggio

Per garantire una governance efficace, il Tavolo tecnico di cui al precedente paragrafo 2.2.2. utilizza indicatori quantitativi e qualitativi, idonei ad offrire scenari di benchmark in scenari pluriennali. Tra questi figurano il numero di strumenti IA in uso, la tipologia di applicazioni e processi coinvolti, le segnalazioni di anomalie e incidenti, gli esiti delle valutazioni di impatto preventive (DPIA), il livello di formazione del personale e il grado di soddisfazione degli utenti interni ed esterni. Tali indicatori permettono di valutare l'efficacia delle misure adottate, l'affidabilità dei sistemi e l'aderenza ai principi di trasparenza, equità e sicurezza, supportando decisioni strategiche e interventi correttivi mirati.

2.6.2 Revisione periodica

Le presenti Linee guida sono soggette a revisione periodica con cadenza almeno biennale, oppure in caso di significativi mutamenti normativi, tecnologici o organizzativi che possano influire sull'uso dei sistemi IA. Il Tavolo tecnico di cui al precedente paragrafo 2.2.2 svolge un ruolo chiave nell'organizzare la raccolta di feedback dagli uffici e dagli utenti, nell'elaborazione di raccomandazioni operative e nella pubblicazione di report periodici. Questo processo garantisce un aggiornamento costante delle procedure, favorendo un'adozione responsabile e sostenibile dell'IA, coerente con i principi etici, giuridici e organizzativi stabiliti a livello regionale.

Glossario

Intelligenza artificiale (IA): Sistemi automatizzati che elaborano dati per generare output (previsioni, raccomandazioni, decisioni, contenuti).

Bias algoritmico: Distorsione sistematica nei risultati generati dall'IA, dovuta a difetti nei dati di addestramento o nella progettazione.

Accountability: Responsabilità giuridica e organizzativa delle decisioni IA.

Human-in-the-loop: Supervisione e validazione umana delle decisioni IA.

Registro regionale degli strumenti IA: Elenco ufficiale delle soluzioni IA autorizzate all'uso nell'Amministrazione regionale.

DPIA (Data Protection Impact Assessment): Valutazione d'impatto sulla protezione dei dati personali.

GenAI: Sistemi di IA generativa, in grado di creare contenuti nuovi (testi, immagini, audio, codice).

Chatbot: Software che simula una conversazione con un essere umano, spesso utilizzato per fornire assistenza o informazioni.

Tavolo tecnico regionale IA: Organo collegiale incaricato di monitorare, valutare e aggiornare le policy IA della Regione.

A

„Leitlinien für den Einsatz künstlicher Intelligenz bei der
Autonomen Region Trentino-Südtirol“

Inhaltsverzeichnis

Einführung	3
1 Der europäische und nationale Rechtsrahmen	4
1.1 Grundgedanken und Zielsetzungen der europäischen und nationalen Gesetzesbestimmungen	6
1.1.1 Fairness und Nichtdiskriminierung	6
1.1.2 Sicherheit und Datenschutz.....	7
1.1.3 Verhältnismäßigkeit.....	7
1.1.5 Transparenz und Rückverfolgbarkeit	8
1.1.6 Genauigkeit, Überprüfung und Validierung	8
1.1.7 Ethische Grundsätze und Nichtdiskriminierung.....	8
1.1.8 Schutz der Arbeitnehmenden	9
2. Die „Leitlinien für den Einsatz künstlicher Intelligenz bei der Autonomen Region Trentino-Südtirol“	9
2.1 Ziele	9
2.2 Anwendungsbereich und Governance	9
2.2.1 Zielgruppen.....	9
2.2.2 Regionale Arbeitsgruppe für künstliche Intelligenz	10
2.2.3 Ausgeschlossene Technologien und Anwendungen.....	10
2.2.4 Regelung der Ausnahmen	11
2.3 Register, Vorgehensweise und Verantwortlichkeiten	11
2.3.1 Regionales Register der KI-Tools	11
2.4.1 Zulässige Nutzung	14
2.4.2 Unzulässige Nutzung.....	16
2.5. Pflichtschulung.....	17
2.6. Überwachung und Überprüfung	18
2.6.1 Indikatoren für die Überwachung.....	18
2.6.2 Regelmäßige Überprüfung	18
Glossar	19

Einführung

Künstliche Intelligenz (KI) ist heute eine Technologie mit hohem Transformationspotenzial, die Entscheidungs-, Produktions- und Beziehungsprozesse innerhalb der öffentlichen Verwaltung tiefgreifend beeinflussen kann. Ihr Einsatz ist nicht nur eine technologische Chance, sondern auch ein strategisches Instrument, um die Effizienz der Dienste zu verbessern, Verwaltungsabläufe zu vereinfachen und eine unmittelbarere und transparentere Interaktion zwischen Verwaltung, Bürgern und Bürgerinnen sowie Unternehmen zu fördern.

Die Einführung von KI in den Arbeitskontext muss verantwortungsbewusst und reguliert erfolgen. Es ist unerlässlich, dass jedes KI-basierte System unter Einhaltung der Grundsätze der Transparenz, der Rückverfolgbarkeit und des Schutzes der Rechte der Arbeitnehmenden implementiert wird und gleichzeitig einen fairen und inklusiven Zugang zu den neuen Technologien gewährleistet.

Die verantwortungsvolle Einführung von KI erfordert daher einen strukturierten Ansatz, der ethische, verfahrenstechnische und organisatorische Bewertungen umfasst, damit die Auswirkungen auf die Arbeit und auf die internen Abläufe in den Ämtern sowie die möglichen Auswirkungen im institutionellen und sozialen Kontext positiv und nachhaltig sind.

In dieser Hinsicht ist KI kein Selbstzweck, sondern ein Instrument im Dienste der organisatorischen Innovation und der Qualität öffentlicher Dienste. Ihre Einführung muss mit klaren Governance-Maßnahmen, der Schulung des Personals und einer kontinuierlichen Überwachung der automatisierten Entscheidungsprozesse in Übereinstimmung mit den geltenden nationalen und europäischen Rechtsvorschriften einhergehen.

In diesem Sinne verändert die künstliche Intelligenz bereits jetzt die Arbeitswelt grundlegend und wirkt sich unterschiedlich auf die verschiedenen Wirtschafts- und

Produktionsbereiche aus. In Verwaltungsprozessen ermöglicht KI die Automatisierung sich wiederholender Tätigkeiten, die Optimierung des Ressourceneinsatzes und die Unterstützung strategischer Entscheidungen durch fortschrittliche Datenanalyse und Trendprognosen, wo dies im Hinblick auf die Organisationsziele sinnvoll und notwendig ist. KI-basierte Technologien wie maschinelle Lernsysteme, Chatbots, virtuelle Assistenten und Plattformen für prädiktive Analysen tragen zur Steigerung der betrieblichen Effizienz bei und ermöglichen es Organisationen, komplexe Herausforderungen mit innovativen Lösungen zu bewältigen.

1 Der europäische und nationale Rechtsrahmen

Auf europäischer Ebene wird der Rechtsrahmen hauptsächlich durch die **Verordnung (EU) 2024/1689**, den so genannten „AI Act“, festgelegt, die ein einheitliches Rechtssystem für das gesamte Gebiet der Union schafft. Diese Verordnung führt einen risikobasierten Ansatz ein, bei dem KI-Systeme in die vier Risikokategorien „inakzeptabel“, „hoch“, „begrenzt“ und „minimal“ eingeteilt werden. Systeme mit inakzeptablem Risiko, die z. B. Social Scoring oder Verhaltensmanipulation ermöglichen, sind verboten, da sie mit den Grundwerten der Union unvereinbar sind. Systeme mit hohem Risiko, zu denen Anwendungen in sensiblen Bereichen wie Sicherheit und Personalmanagement gehören, sind nur zulässig, wenn sie strenge Anforderungen an Sicherheit, Transparenz und menschlicher Kontrolle erfüllen. Systeme mit begrenztem Risiko, wie Chatbots und Tools zur Generierung synthetischer Inhalte oder Deep Fakes, unterliegen besonderen Transparenzpflichten gegenüber Nutzern und Nutzerinnen, während Systeme mit minimalem Risiko unter Einhaltung der allgemeinen Grundsätze der Verordnung frei verwendet werden können.

Zu den wichtigsten Vorschriften des „AI Act“ gehören die Bewertung der Auswirkungen auf die Grundrechte (sog. Grundrechte-Folgenabschätzung) und der mit dem Einsatz von KI verbundenen Systemrisiken, die Führung eines Registers der eingesetzten KI-Tools, die Verpflichtung zur Gewährleistung einer wirksamen menschlichen Aufsicht in allen Phasen des automatisierten Entscheidungsprozesses und die Transparenz der Informationen gegenüber den Nutzern und Nutzerinnen. Diese Maßnahmen sollen die Rückverfolgbarkeit von Entscheidungen, die Überprüfbarkeit von Ergebnissen und die

Verhinderung diskriminierender oder die individuellen Rechte verletzender Auswirkungen gewährleisten.

Parallel dazu legt die **Verordnung (EU) 2016/679**, allgemein bekannt als „Datenschutz-Grundverordnung“ (DSGVO), den Rahmen für den Schutz personenbezogener Daten fest und führt die Grundsätze der Rechtmäßigkeit, Korrektheit, Transparenz, Datenminimierung, Speicherbegrenzung, Integrität und Vertraulichkeit ein. Die DSGVO erkennt den betroffenen Personen Grundrechte zu, darunter das Recht auf Zugang, Berichtigung, Löschung, Einschränkung der Verarbeitung, Widerspruch und Übertragbarkeit ihrer Daten. Der Datenschutzbeauftragte (Data Protection Officer – DPO) nimmt eine zentrale Rolle ein, da er für die Beratung, die Überwachung der Anwendung der Vorschriften sowie die Förderung interner Schulungsmaßnahmen zuständig ist und auch als Ansprechpartner für die Datenschutzbehörde fungiert.

Auf nationaler Ebene sind im „Kodex der digitalen Verwaltung“ (**GvD Nr. 82/2005**) die digitalen Rechte der Bürger und Bürgerinnen und die Pflichten der öffentlichen Verwaltungen in Bezug auf die Verwaltung von Online-Diensten, digitale Identität, Interoperabilität der Systeme, Sicherheit der Infrastrukturen, Aufbewahrung und Zugänglichkeit von elektronischen Dokumenten festgelegt. In engem Zusammenhang damit schreiben die Bestimmungen über Transparenz und Korruptionsvorbeugung (GvD Nr. 33/2013) Verpflichtungen zur Bekanntmachung und Rückverfolgbarkeit von Verwaltungstätigkeiten vor, deren Ziel es ist, eine KI-Nutzung zu gewährleisten, die den Grundsätzen der Unparteilichkeit und der guten Führung der öffentlichen Verwaltung entspricht.

Eine wichtige Rolle spielen auch die Leitlinien und technischen Dokumente, die von nationalen und europäischen Einrichtungen ausgearbeitet werden. Die Agenzia per l'Italia Digitale (AgID) legt Standards und Empfehlungen in den Bereichen Interoperabilität, Sicherheit, Einführung von Cloud-Systemen und Dokumentenverwaltung fest. Die Agentur der Europäischen Union für Cybersicherheit (ENISA) erteilt praktische Hinweise zum Umgang mit IT-Risiken und zur Sicherheit digitaler Systeme. Die Datenschutzbehörde

veröffentlicht Stellungnahmen, Empfehlungen und Sanktionsmaßnahmen, um die Einhaltung der Vorschriften zum Datenschutz und zur Datenverarbeitung sicherzustellen.

Schließlich führt das kürzlich verabschiedete **Gesetz Nr. 132/2025** (in Kraft getreten am 10. Oktober 2025) im Artikel 14 („Einsatz künstlicher Intelligenz in der öffentlichen Verwaltung“) die Verpflichtung ein, ein Governance- und Transparenzsystem für den Einsatz von KI in öffentlichen Körperschaften einzurichten.

1.1 Grundgedanken und Zielsetzungen der europäischen und nationalen Gesetzesbestimmungen

Der gesamte Regulierungsrahmen basiert auf den Verfassungsgrundsätzen der Rechtmäßigkeit, Unparteilichkeit, guten Führung und Transparenz, die die Grundlagen der Verwaltungstätigkeit darstellen und die unverzichtbare Wertebasis für die Anwendung künstlicher Intelligenz im öffentlichen Bereich bilden. Diese Grundsätze gewährleisten, dass technologische Innovation die demokratischen Garantien nicht beeinträchtigt, sondern sich als Instrument zur Unterstützung der Gemeinschaft unter Achtung der Würde und der Rechte der Menschen entwickelt. Insbesondere sind die Regionalverwaltung, ihre Bediensteten und die im Namen und im Auftrag dieser Verwaltung tätigen Personen bei der Nutzung von KI zur Einhaltung der folgenden Grundsätze verpflichtet.

1.1.1 Fairness und Nichtdiskriminierung

Die Einführung von Systemen der künstlichen Intelligenz muss auf den Grundsätzen der Fairness, Unparteilichkeit und Gleichbehandlung beruhen. Jede Form von algorithmischen Verzerrungen, die zu direkter oder indirekter Diskriminierung aufgrund von Geschlecht, ethnischer Zugehörigkeit, Alter, Orientierung, sozioökonomischem Status oder anderen persönlichen Merkmalen führen könnte, muss verhindert und korrigiert werden. Zu diesem Zweck müssen Prozesse zur kontinuierlichen Überwachung der von den Systemen generierten Ergebnisse implementiert werden, um etwaige systematische Abweichungen zu erkennen und Korrekturmaßnahmen zu ergreifen. Fairness muss auch in der Phase der Modellentwicklung und des Modelltrainings durch die Auswahl repräsentativer Datensätze und die unabhängige Validierung der Algorithmen gewährleistet werden. Die Transparenz der Bewertungskriterien und die Möglichkeit einer menschlichen Überprüfung der Ergebnisse sind weitere Garantien für Unparteilichkeit.

1.1.2 Sicherheit und Datenschutz

Die IT-Sicherheit und der Schutz personenbezogener Daten sind wesentliche Voraussetzungen für den Einsatz künstlicher Intelligenz. Es ist verboten, sensible, klassifizierte oder anderweitig zur Offenlegung vertraulicher Informationen geeignete Daten in KI-Systeme einzugeben, es sei denn, es liegen spezielle Genehmigungen vor und es wurden angemessene Sicherheitsmaßnahmen wie Verschlüsselung, Pseudonymisierung oder kontrollierter Zugriff getroffen. Die Verarbeitung muss den Grundsätzen der DSGVO und den Leitlinien der Datenschutzbehörde entsprechen und es muss sichergestellt werden, dass die Daten für bestimmte, relevante und auf den institutionellen Zweck beschränkte Zwecke erhoben werden. Parallel zum Einsatz von KI-Tools müssen außerdem Maßnahmen zur Daten-Governance umgesetzt werden, die die Integrität, Verfügbarkeit und Rückverfolgbarkeit der Daten gewährleisten und Pläne für die Reaktion auf Vorfälle (Incident Response) und regelmäßige Audits umfassen, um Datenschutzverletzungen oder unerlaubte Zugriffe zu verhindern.

1.1.3 Verhältnismäßigkeit

Der Grundsatz der Verhältnismäßigkeit erfordert, dass der Einsatz künstlicher Intelligenz eng auf die institutionellen Ziele und die Zwecke der Datenverarbeitung abgestimmt sein muss. Ein übermäßiger oder überflüssiger Einsatz der Technologie sowie die Erhebung von Informationen, die im Hinblick auf die verfolgten Ziele unnötig oder unverhältnismäßig sind, sind zu vermeiden. Die Auswirkungen der KI müssen nicht nur im Hinblick auf die Effizienz bewertet werden, sondern auch im Hinblick auf die potenziellen Auswirkungen auf die Privatsphäre, die Entscheidungsautonomie der Arbeitnehmer und Arbeitnehmerinnen und die Qualität der den Bürgern und Bürgerinnen angebotenen Dienstleistungen. Verhältnismäßigkeit bedeutet daher ein dynamisches Gleichgewicht zwischen Innovation und Schutz, das sicherstellt, dass die Automatisierung die menschlichen Fähigkeiten nicht ersetzt, sondern unterstützt.

1.1.4 Ethische Nutzung und Zentralität des Menschen

KI-Systeme müssen als Hilfsmittel betrachtet werden, die mit dem Ziel entwickelt wurden, das Gemeinwohl und die Qualität öffentlicher Dienstleistungen zu verbessern, ohne die geistigen, kreativen und Entscheidungsfähigkeiten des Menschen zu ersetzen. Die menschliche Aufsicht bleibt ein unverzichtbares Prinzip: Der Nutzer/Die Nutzerin ist stets

verantwortlich für den Prozess und für die endgültige Entscheidung, die nicht vollständig an die Maschine delegiert werden kann. KI erfüllt in diesem Zusammenhang eine unterstützende und verstärkende Funktion für die menschlichen Fähigkeiten und trägt zur Optimierung der Arbeitseffizienz und der Verwaltung der Informationen bei, ohne jedoch Entscheidungsautonomie in sensiblen oder ermessensabhängigen Bereichen zu übernehmen.

1.1.5 Transparenz und Rückverfolgbarkeit

Der Einsatz von KI-Tools bei Arbeitstätigkeiten muss transparent und rückverfolgbar sein. Es muss klar angegeben werden, wenn ein Output, sei es ein Dokument, ein Schreiben oder eine Antwort auf Anfragen, durch ein KI-System generiert oder ergänzt wurde. Dieser Grundsatz ermöglicht es, das Vertrauen in die erzielten Ergebnisse zu bewahren und den Algorithmus als Arbeitsinstrument und nicht als Urheber der Informationen zu erkennen.

1.1.6 Genauigkeit, Überprüfung und Validierung

Jeder Output eines KI-Systems muss einem strengen Überprüfungs- und Validierungsprozess unterzogen werden, um die Genauigkeit, Vollständigkeit und Relevanz der Informationen zu gewährleisten. Eine kritische Bewertung ist unerlässlich, da KI-Systeme unvollständige, veraltete oder in einigen Fällen sogar falsche oder irreführende Informationen, sogenannte „Halluzinationen“, generieren können. Um diese Risiken zu mindern, müssen Arbeitsinstrumente wie Validierungsschecklisten für jeden Output und Gegenprüfungen durch mindestens zwei Mitarbeiter bei Dokumenten mit größerer Auswirkung eingesetzt werden. Dieser Ansatz gewährleistet eine systematische Qualitätskontrolle, verringert die Fehlerwahrscheinlichkeit und erhöht das Vertrauen in den Einsatz von KI in Entscheidungs- und Arbeitsprozessen.

1.1.7 Ethische Grundsätze und Nichtdiskriminierung

KI-Systeme müssen so konzipiert und eingesetzt werden, dass Verzerrungen und Diskriminierung vermieden werden, grundlegende ethische Prinzipien eingehalten werden und Inklusion aktiv gefördert wird. Ziel ist eine faire und unparteiische Behandlung aller Personen und Gruppen, wobei den am stärksten benachteiligten Gruppen besondere Aufmerksamkeit zu widmen ist. Etwaige Bias in den Daten oder Algorithmen müssen überwacht und korrigiert werden, indem Strategien zur Minderung der Auswirkungen von Verzerrungen ergriffen werden, um sicherzustellen, dass die erzielten Ergebnisse fair und

transparent sind und den ethischen und normativen Werten der öffentlichen Verwaltung entsprechen.

1.1.8 Schutz der Arbeitnehmenden

Die Einführung von KI-Systemen am Arbeitsplatz muss unter Einhaltung des sozialen Dialogs und der arbeitsrechtlichen Bestimmungen erfolgen. Der Einsatz von künstlicher Intelligenz zur Überwachung der Tätigkeiten der Arbeitnehmer und Arbeitnehmerinnen erfordert klare Regeln, um invasive Formen der Überwachung oder übermäßige Kontrollen zu vermeiden, die die Freiheit, die Würde und die Qualität der Arbeit beeinträchtigen könnten. Die Technologie muss die organisatorische Effizienz und die Sicherheit der Prozesse unterstützen, ohne die Arbeitsbedingungen negativ zu beeinflussen, und dabei ein Gleichgewicht zwischen Innovation, Produktivität und Schutz der Rechte der Arbeitnehmenden gewährleisten.

2. Die „Leitlinien für den Einsatz künstlicher Intelligenz bei der Autonomen Region Trentino-Südtirol“

2.1 Ziele

Im Rahmen der im vorstehenden Kapitel 1 dargelegten allgemeinen Grundsätze erkennt die Autonome Region Trentino-Südtirol, auch aufgrund der ihr durch das Sonderstatut übertragenen Befugnisse und im Rahmen ihrer territorialen Zuständigkeiten, die Notwendigkeit an, im Zusammenhang mit und in Ergänzung zu den bereits erwähnten europäischen und nationalen Rechtsvorschriften einen umfassenden Rahmen mit einheitlichen und verbindlichen Richtlinien und Regeln (Leitlinien) für die Einführung, Entwicklung und Verwaltung von KI-Systemen in den verschiedenen Verwaltungsbereichen zu definieren, um eine unkontrollierte oder uneinheitliche Nutzung der Technologie zu verhindern. Diese Initiative ist Teil des Prozesses der digitalen Transformation des öffentlichen Bereichs und trägt der Notwendigkeit Rechnung, eine Verwendung der KI zu gewährleisten, die den allgemeinen für sie geltenden Grundsätzen entspricht.

2.2 Anwendungsbereich und Governance

2.2.1 Zielgruppen

Die Leitlinien gelten für das gesamte Personal der Autonomen Region Trentino-Südtirol, einschließlich externer Mitarbeiter/Mitarbeiterinnen, Berater/Beraterinnen und Personen,

die in irgendeiner Weise KI-Systeme im Auftrag der Region oder deren Organisationsstrukturen nutzen. Der Anwendungsbereich erstreckt sich – im Rahmen ihrer jeweiligen Zuständigkeiten und Funktionen – auch auf Gesellschaften, Körperschaften und Einrichtungen, die von der Regionalverwaltung kontrolliert oder koordiniert werden.

Ebenfalls in den Anwendungsbereich dieser Leitlinien fallen alle Tätigkeiten, die die Nutzung von Konten, digitalen Tools oder Datenbanken beinhalten, die im Rahmen der Arbeitsaufgaben erworben oder verwaltet werden, unabhängig von der Art des verwendeten Geräts und davon, ob es von der Verwaltung bereitgestellt wird oder Eigentum des Mitarbeiters / der Mitarbeiterin ist. Dies umfasst jede Form der Interaktion mit KI-Systemen, die der Datenverarbeitung, der Erstellung von Inhalten oder der Entscheidungsunterstützung im institutionellen Bereich dient.

2.2.2 Technische Arbeitsgruppe der Region für künstliche Intelligenz

Es wird die „Technische Arbeitsgruppe der Region für Künstliche Intelligenz“ für die Governance, Koordinierung, Einführung, Implementierung und Überwachung von KI-Systemen eingerichtet. Die Zusammensetzung der Arbeitsgruppe wird durch Beschluss der Regionalregierung festgelegt und muss die Führungskräfte der Region, Bedienstete des Amtes für Informatik und Digitalisierung sowie den Datenschutzbeauftragten (DPO) umfassen. Je nach den zu behandelnden Themen kann die Arbeitsgruppe von Fall zu Fall durch weitere fachkundige Mitarbeiter ergänzt werden. Die Technische Arbeitsgruppe ist mit der Aktualisierung der Regeln und Verfahren entsprechend der rechtlichen und technologischen Entwicklung, einschließlich der vorliegenden Leitlinien, sowie mit der Bewertung von Meldungen und der Bearbeitung von Anomalien, einschließlich vorübergehender Aussetzungen, interner Audits oder Überarbeitungen der Richtlinien, beauftragt.

Die Technische Arbeitsgruppe übt außerdem die im folgenden Abschnitt 2.2.4. genannten Zuständigkeiten aus.

2.2.3 Ausgeschlossene Technologien und Anwendungen

Aus dem Anwendungsbereich dieser Leitlinien ausgeschlossen sind IT-Technologien, die keine auf künstlicher Intelligenz basierenden Merkmale aufweisen, d. h. Anwendungen, die nach deterministischen Logiken arbeiten und nicht auf autonomem Lernen oder autonomer Anpassung basieren. Zu den ausgeschlossenen Anwendungen gehören auch gewöhnliche Automatisierungen ohne autonome Lern- oder Entscheidungsfähigkeiten, wie

beispielsweise statische Workflow-Systeme, Makro-Anwendungen oder Dokumentenmanagement-Software, die keine prädiktiven oder generativen Modelle enthalten.

Diese Ausschlüsse entbinden die Nutzer und Nutzerinnen jedoch nicht von der Verpflichtung, die Sicherheit, Vertraulichkeit und ordnungsgemäße Verwaltung der verarbeiteten Daten zu gewährleisten, die weiterhin den geltenden Vorschriften zum Datenschutz und zur guten Führung der Verwaltung unterliegen.

2.2.4 Regelung der Ausnahmen

Eventuelle Ausnahmen von der Anwendung dieser Leitlinien können nur in begründeten und begrenzten Fällen nach Bewertung und Genehmigung durch die im Abschnitt 2.2.2 genannte Technische Arbeitsgruppe zugelassen werden. Dem Antrag auf Ausnahme sind detaillierte Unterlagen beizulegen, aus denen die Gründe für die Ausnahme, die Art des betreffenden Systems oder Projekts sowie die gegebenenfalls getroffenen Maßnahmen zur Risikominderung hervorgehen, mit denen die Einhaltung der allgemeinen Grundsätze der Rechtmäßigkeit, Transparenz und Sicherheit gewährleistet werden soll.

Die Technische Arbeitsgruppe hat die Aufgabe, die Anträge zu prüfen, die Übereinstimmung mit dem Rechtsrahmen zu überprüfen und die potenziellen Auswirkungen auf die Organisation und die Rechte der Nutzer und Nutzerinnen zu bewerten. Die getroffenen Entscheidungen müssen zu Kontroll- und Überprüfungs Zwecken protokolliert und aufbewahrt werden, um sicherzustellen, dass Ausnahmen auch Ausnahmen bleiben und nicht zur gängigen Praxis werden.

Auf diese Weise sorgt die Regelung der Ausnahmen für ein Gleichgewicht zwischen der Notwendigkeit von Flexibilität in der Verwaltung und der Notwendigkeit der Aufrechterhaltung hoher Standards in Bezug auf Zuverlässigkeit, Sicherheit und Verantwortung bei der Nutzung künstlicher Intelligenz bei der Region.

2.3 Register, Vorgehensweise und Verantwortlichkeiten

2.3.1 Regionales Register der KI-Tools

Beim Amt für Informatik und Digitalisierung wird das „Regionale Register der KI-Tools“ eingerichtet, das gemäß den Richtlinien der im Abschnitt 2.2.2 genannten Technischen Arbeitsgruppe arbeitet, um die technische, rechtliche und organisatorische Konformität der eingesetzten Lösungen zu gewährleisten.

Das Register enthält eine aktualisierte Liste der zugelassenen KI-Tools, gegebenenfalls ergänzt durch technische Datenblätter, in denen die Hauptmerkmale, Verwendungszwecke und Anwendungsbereiche beschrieben werden. Die Datenblätter enthalten auch Informationen zur Folgenabschätzung, zu den implementierten Sicherheitsmaßnahmen, zum vorgesehenen Umfang der menschlichen Aufsicht und zu eventuellen operativen Einschränkungen. Jedes registrierte Tool wird hinsichtlich seines Aktualisierungsstatus, der durchgeführten Überprüfungen und etwaiger Meldungen oder Aussetzungen, die seine Zuverlässigkeit oder Konformität mit den Vorschriften beeinträchtigen könnten, überwacht. Die Aufnahme eines neuen KI-Tools in das Register unterliegt einem dreistufigen Bewertungsverfahren:

1. Die technische Bewertung: Sie betrifft die IT-Sicherheit, die Interoperabilität mit bestehenden Systemen und die operative Sicherheit des KI-Tools, und gewährleistet, dass die Anwendung ohne Beeinträchtigung der Datenintegrität oder der Prozesseffizienz eingesetzt werden kann.
2. Die rechtliche Bewertung: Es wird die Übereinstimmung des KI-Tools mit den geltenden Vorschriften überprüft, einschließlich der Notwendigkeit einer Datenschutz-Folgenabschätzung (DSFA) und der Einhaltung der Anforderungen des AI Act, der DSGVO und anderer einschlägiger Vorschriften.
3. Die organisatorische Bewertung: Es werden die Auswirkungen der Einführung des KI-Tools auf die internen Prozesse analysiert und die für eine sichere und effektive Nutzung erforderlichen Verfahren, operativen Änderungen und Schulungen des Personals ermittelt.

2.3.2 Verhaltensrichtlinien für das Personal und erste zulässige KI-Lösung

Es wird erneut bekräftigt, dass alle Bediensteten, Mitarbeiter/Mitarbeiterinnen oder Berater/Beraterinnen, die im Auftrag der Autonomen Region Trentino-Südtirol KI-Systeme nutzen, verpflichtet sind, nach den Grundsätzen der Sorgfalt, Verantwortung und Transparenz zu handeln und sicherzustellen, dass die Nutzung dieser Tools in voller Übereinstimmung mit den geltenden Vorschriften und internen organisatorischen Bestimmungen erfolgt. Der Einsatz von KI muss stets von einer aktiven menschlichen Aufsicht begleitet werden, um die Richtigkeit, Zuverlässigkeit und Relevanz der

generierten Ergebnisse zu überprüfen. Jedes von einem KI-System erzeugte oder unterstützte Ergebnis muss vom verantwortlichen Nutzer validiert werden, der die volle Verantwortung in technischer, rechtlicher und inhaltlicher Hinsicht übernimmt.

Insbesondere ist das Personal verpflichtet, in allen offiziellen Dokumenten, Berichten, internen Mitteilungen oder Verwaltungsakten ausdrücklich darauf hinzuweisen, wenn der Inhalt mit Hilfe eines KI-Systems erstellt, geändert oder ergänzt wurde, indem am Ende des Dokuments der folgende oder ein ähnlicher Satz angefügt wird: „*Auch mit Hilfe von KI erstellt*“. Diese Transparenzpflicht entspricht dem Grundsatz der Rückverfolgbarkeit von Entscheidungsprozessen und gewährleistet die Möglichkeit, die Rolle der KI bei der Erstellung des Dokuments in Übereinstimmung mit den Anforderungen der Rechenschaftspflicht und der öffentlichen Kontrolle nachzuvollziehen. Es ist nicht erforderlich, die Verwendung von KI für interne Tätigkeiten oder für die Erstellung von vorläufigen Entwürfen anzugeben, die dann vom Nutzer / von der Nutzerin erheblich geändert werden, wobei natürlich die endgültige Überprüfung und Validierung durch diesen unbeschadet bleibt.

Es ist verboten, ermessensabhängige Entscheidungen vollständig an KI-Systeme zu delegieren, insbesondere in sensiblen Bereichen oder in Bereichen, die direkte Auswirkungen auf die Rechte und Interessen von Personen haben können. Dazu gehören beispielsweise Entscheidungen über die Gewährung von finanziellen Vorteilen, die Verhängung von Strafen, die Durchführung von Disziplinarverfahren oder die Bewertung der Leistung des Personals. In solchen Fällen kann die KI eine rein informative oder analytische Unterstützungsfunktion übernehmen, während die endgültige Entscheidung ausschließlich beim Menschen liegen muss.

Da die Verwendung von KI-Systemen nur zulässig ist, wenn diese im Regionalen Register der KI-Tools eingetragen sind, muss das Personal vorab überprüfen, ob die verwendete Anwendung zugelassen und auf dem neuesten Stand ist sowie den Anforderungen an Sicherheit, Datenschutz und Interoperabilität entspricht. Die Verwendung nicht registrierter oder nicht zugelassener Anwendungen stellt einen Verstoß gegen diese Leitlinien dar und kann Disziplinar- oder Verwaltungsmaßnahmen nach sich ziehen.

Folgende KI-Tools, die in das im Abschnitt 2.3.1 genannte Register aufgenommen wurden und bereits auf den E-Mail-Konten der Verwaltung (mit der Domain @regione.taa.it) aktiv sind, sind zum Zeitpunkt der Verabschiedung dieser Leitlinien zugelassen:

- **Gemini IA:** Es handelt sich um einen KI-Assistenten, der die Gemini-Sprachmodelle von Google nutzt. Er wurde entwickelt, um die Kreativität und Produktivität der Nutzer und Nutzerinnen zu steigern, und bietet eine Vielzahl von Funktionen.
- **Gemini Workspace:** Hierbei handelt es sich um die künstliche Intelligenz von Gemini, die direkt in die Google-Apps für Produktivität und Zusammenarbeit integriert ist, darunter Gmail, Docs, Sheets, Präsentationen und Drive.

Es dürfen ausschließlich die gemäß diesen Leitlinien ausdrücklich genehmigten KI-Tools (und GenAI-Tools) verwendet werden. Auch die Verwendung von KI-basierten (oder GenAI-basierten) Tools, die nicht gemäß diesen Leitlinien genehmigt sind, über persönliche Konten für berufliche Zwecke ist nicht gestattet.

2.4 Praktische Einsatzbeispiele und Anwendungsfälle

2.4.1 Zulässige Nutzung

Der Einsatz von KI-Systemen bei der Autonomen Region Trentino-Südtirol bringt eine Reihe von Vorteilen mit sich, die sich auf die Qualität der Arbeit, den Zugang zu Informationen, die Produktivität, die Kreativität, die Zusammenarbeit, die Qualität der Entscheidungen, die Sicherheit und die Nachhaltigkeit der Verwaltungstätigkeit auswirken. Allerdings müssen stets eine menschliche Aufsicht (Human-in-the-Loop) und die volle Verantwortung des Nutzers / der Nutzerin gewährleistet sein. Die Nutzung muss den Grundsätzen der Rechtmäßigkeit, Transparenz und Nichtdiskriminierung sowie den Bestimmungen laut Abschnitt 1.1 dieser Leitlinien entsprechen.

Zu den zulässigen Anwendungsbereichen gehören:

- A. Arbeitsunterstützung und Dokumentenanalyse

KI kann zur Optimierung der internen Arbeitsabläufe eingesetzt werden, um sich wiederholende Tätigkeiten zu reduzieren. Beispielsweise kann KI folgende Rolle/Funktion übernehmen:

- Assistent für die Erstellung nicht vertraulicher Dokumente: Erstellung komplexer Dokumente, Diskussionen oder elektronischer Mitteilungen, Suche nach Quellen in großen Datenmengen, Brainstorming und Ideenfindung für die Ausarbeitung der endgültigen Fassung.
- Erstellung von Entwürfen und Inhalten: Unterstützung bei der Erstellung von ersten Entwürfen für Texte, Berichte, interne Mitteilungen oder nicht bindende Dokumente, sofern die Ergebnisse einer strengen Überprüfung und Validierung durch das Personal unterzogen werden.
- Klassifizierung und Sortierung von Dokumenten: Einsatz von KI-Systemen zur automatischen Klassifizierung und Sortierung eingehender Dokumente, mit kontinuierlicher Überwachung von Anomalien und spezifischer Schulung des zuständigen Personals.

B. Verbesserung der Dienstleistungen für die Bürger und Bürgerinnen

KI ermöglicht besser zugängliche und zeitnahe Antworten an die Öffentlichkeit.

- Virtueller Assistent (Chatbot): Implementierung von Chatbots und virtuellen Assistenten, um häufig gestellte Fragen (FAQ) und allgemeine Anfragen von Bürgern und internen Mitarbeitern direkt zu beantworten. Bei komplexen Antworten oder offiziellen Veröffentlichungen ist eine vorherige menschliche Validierung erforderlich.
- Zugänglichkeit und Inklusion: Mehrsprachige Sprachunterstützung und Tools, die kommunikative oder physische Barrieren abbauen und einen gerechteren und besser nutzbaren öffentlichen Dienst gewährleisten.

C. Unterstützung bei Entscheidungs- und Strategieprozessen

KI kann erweiterte Analysen liefern und so die Qualität menschlicher Entscheidungen verbessern.

- Prädiktive Analyse und Simulationen: Verarbeitung großer Datenmengen zur Erstellung von Schätzungen, Zukunftsszenarien, Advanced Reporting und

Simulationen, die strategische und verwaltungstechnische Entscheidungen unterstützen.

- Vorab-Analyse für Beiträge/Ausschreibungen: Einsatz von Systemen zur Vorab-Analyse von Anträgen auf Gewährung wirtschaftlicher Vorteile oder Finanzierungen, sofern die endgültige Entscheidung stets von einem Beamten / einer Beamtin getroffen wird. Es müssen regelmäßige Audits der Ergebnisse vorgesehen werden, um Verzerrungen zu vermeiden.

In jedem Fall ist es Pflicht:

1. vorab zu überprüfen, ob das Tool im Regionalen Register der KI-Tools eingetragen und zugelassen ist;
2. in alle offiziellen Akten und Dokumente eine ausdrückliche Erklärung über die Verwendung von künstlicher Intelligenz aufzunehmen, wie im Abschnitt 2.3.2, zweiter Absatz, angegeben;
3. vor der Implementierung von Systemen, die personenbezogene Daten verarbeiten, eine Datenschutz-Folgenabschätzung (Data Protection Impact Assessment, DSFA) durchzuführen, wie von den einschlägigen Bestimmungen vorgeschrieben.

2.4.2 Unzulässige Nutzung

Der Einsatz von KI-Systemen bei der Autonomen Region Trentino-Südtirol unterliegt strengen Auflagen zum Schutz der Sicherheit, der Privatsphäre, der Ethik und der Einhaltung gesetzlicher Vorschriften. Es ist verboten, personenbezogene Daten oder als vertraulich eingestufte Informationen in die Systeme einzugeben, um Risiken einer unzulässigen Offenlegung, eines unbefugten Zugriffs oder von Verstößen gegen Datenschutzbestimmungen vorzubeugen.

Es ist nicht gestattet, KI-Systeme zu verwenden, um diskriminierende, beleidigende, gewalttätige, illegale oder die Rechte Dritter verletzende Inhalte zu generieren oder zu verbreiten. Ebenso ist es verboten, die Ergebnisse eines KI-Systems als Originalarbeit zu präsentieren, ohne dessen Verwendung anzugeben, da dies einen Verstoß gegen die Grundsätze der Transparenz und Integrität darstellt.

Die Verwendung von KI-Systemen zur automatischen Entscheidungsfindung mit erheblichen Auswirkungen auf Personen oder Prozesse ohne menschliche Aufsicht und Validierung ist verboten, da dies die Verantwortlichkeit, die Rückverfolgbarkeit und den

Grundsatz der zentralen Rolle des Menschen in Entscheidungsprozessen beeinträchtigt. Ebenso ist es nicht zulässig, KI zu verwenden, um sich unbefugten Zugang zu Informationen oder Systemen zu verschaffen oder für nicht dienstbezogene oder persönliche Tätigkeiten, die die institutionellen Funktionen der Körperschaft beeinträchtigen könnten.

Darüber hinaus ist es verboten, auf Emotionserkennung, Eye-Tracking oder Stimmungsanalyse basierte Technologien zur Überwachung der Bediensteten einzusetzen, da diese Anwendungen die Privatsphäre, das Wohlbefinden und die Rechte der Beschäftigten beeinträchtigen können. Der Einsatz von KI-Systemen muss stets im Einklang mit den Gesetzen, Vorschriften und internen Richtlinien der Körperschaft stehen und sicherstellen, dass die Anwendungen mit den institutionellen, ethischen und operativen Zielen vereinbar sind.

2.5. Pflichtschulung

Die für die Weiterbildung des Personals zuständige Organisationsstruktur der Region gibt in ihrer Planung die regelmäßigen Pflichtschulungen für alle Bediensteten vor, die KI-Systeme bei ihrer Arbeit nutzen. Diese Schulungsprogramme zielen darauf ab, fachübergreifende Kompetenzen zu entwickeln, indem sie technische, rechtliche und ethische Kenntnisse miteinander verbinden. Die Kurse behandeln das Potenzial und die Grenzen der KI, die für die sichere und korrekte Nutzung der Systeme erforderlichen technischen Kompetenzen sowie die rechtlichen und Compliance-bezogenen Aspekte in Zusammenhang mit dem Datenschutz und der Überwachung automatisierter Prozesse. Besonderes Augenmerk wird auf die ethische Kultur der Innovation, die Erkennung und Minderung von Verzerrungen, die Verhinderung von Diskriminierung, die Transparenz und Rückverfolgbarkeit von Prozessen sowie die Festlegung von Rollen und Verantwortlichkeiten bei der menschlichen Überwachung gelegt. Die Kurse umfassen auch die Analyse konkreter Anwendungsfälle und Arbeitsszenarien, um die erworbenen Kompetenzen zu festigen und eine verantwortungsvolle und bewusste Anwendung von KI-Systemen zu fördern.

2.6. Überwachung und Überprüfung

2.6.1 Indikatoren für die Überwachung

Um eine wirksame Governance zu gewährleisten, verwendet die im Abschnitt 2.2.2 genannte Technische Arbeitsgruppe quantitative und qualitative Indikatoren, die geeignet sind, Benchmark-Szenarien für mehrjährige Zeiträume zu liefern. Dazu gehören die Anzahl der eingesetzten KI-Tools, die Art der betroffenen Anwendungen und Prozesse, die Meldungen von Anomalien und Vorfällen, die Ergebnisse der Datenschutz-Folgenabschätzungen (DSFA), der Stand der Personalschulung und der Grad der Zufriedenheit der internen und externen Nutzer und Nutzerinnen. Anhand dieser Indikatoren lassen sich die Wirksamkeit der getroffenen Maßnahmen, die Zuverlässigkeit der Systeme und die Einhaltung der Grundsätze der Transparenz, Fairness und Sicherheit bewerten, wodurch strategische Entscheidungen und gezielte Korrekturmaßnahmen unterstützt werden.

2.6.2 Regelmäßige Überprüfung

Die vorliegenden Leitlinien werden regelmäßig – mindestens alle zwei Jahre oder bei wesentlichen rechtlichen, technologischen oder organisatorischen Änderungen, die sich auf die Nutzung von KI-Systemen auswirken können – überprüft. Die im Abschnitt 2.2.2 genannte Technische Arbeitsgruppe spielt eine Schlüsselrolle bei der Organisation der Sammlung von Rückmeldungen aus den Ämtern und von den Nutzern und Nutzerinnen, bei der Ausarbeitung von praktischen Empfehlungen und bei der Veröffentlichung regelmäßiger Berichte. Dieser Prozess gewährleistet eine ständige Aktualisierung der Verfahren und fördert eine verantwortungsvolle und nachhaltige Einführung der KI im Einklang mit den auf regionaler Ebene festgelegten ethischen, rechtlichen und organisatorischen Grundsätzen.

Glossar

Künstliche Intelligenz (KI): Automatisierte Systeme, die Daten verarbeiten, um Ergebnisse (Prognosen, Empfehlungen, Entscheidungen, Inhalte) zu generieren

Algorithmischer Bias: Systematische Verzerrung bei den von der KI generierten Ergebnissen aufgrund von Fehlern in den Trainingsdaten oder Projektierungsfehlern

Verantwortlichkeit: Rechtliche und organisatorische Verantwortung für KI-Entscheidungen

Human-in-the-loop: Menschliche Überwachung und Validierung von KI-Entscheidungen

Regionales Register der KI-Tools: Offizielle Liste der KI-Lösungen, die für den Einsatz bei der Regionalverwaltung zugelassen sind

DSFA: Datenschutz-Folgenabschätzung

GenAI: Generative KI-Systeme, die in der Lage sind, neue Inhalte (Texte, Bilder, Audiomaterialien, Codes) zu erstellen

Chatbot: Software, die eine Unterhaltung mit einem Menschen simuliert und häufig zur Bereitstellung von Unterstützung oder Informationen eingesetzt wird

Technische Arbeitsgruppe der Region für künstliche Intelligenz: mit der Überwachung, Bewertung und Aktualisierung der KI-Richtlinien der Region beauftragtes Gremium